



A Lightweight Mobile System for Crop Disease Diagnosis

Siricharoen, P., Scotney, B., Morrow, P., & Parr, G. (2016). A Lightweight Mobile System for Crop Disease Diagnosis. In *Unknown Host Publication* (Vol. 9730, pp. 783-791). Springer. https://doi.org/10.1007/978-3-319-41501-7_87

[Link to publication record in Ulster University Research Portal](#)

Published in:
Unknown Host Publication

Publication Status:
Published (in print/issue): 15/07/2016

DOI:
[10.1007/978-3-319-41501-7_87](https://doi.org/10.1007/978-3-319-41501-7_87)

Document Version
Author Accepted version

General rights
Copyright for the publications made accessible via Ulster University's Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Ulster University's institutional repository that provides access to Ulster's research outputs. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact pure-support@ulster.ac.uk.

A Lightweight Mobile System for Crop Disease Diagnosis

Punnarai Siricharoen, Bryan Scotney, Philip Morrow, and Gerard Parr

School of Computing and Information Engineering, University of Ulster, United Kingdom

`siricharoen-p@email.ulster.ac.uk`

`{bw.scotney, pj.morrow, gp.parr}@ulster.ac.uk`

Abstract. This paper presents a low-complexity mobile application for automatically diagnosing crop diseases in the field. In an initial pre-processing stage, the system leverages the capability of a smartphone device and basic image processing algorithms to obtain consistent leaf orientation and to remove the background. A number of different features are then extracted from the leaf, including texture, colour and shape features. Nine lightweight sub-features are combined and implemented as a feature descriptor for this mobile environment. The system is applied to six wheat leaf types: non-disease, yellow rust, Septoria, brown rust, powdery mildew and tan spots, which are commonly occurring wheat diseases worldwide. The standalone application demonstrates the possibilities for disease diagnosis under realistic circumstances, with disease/non-disease detection accuracy of approximately 88%, and can provide a possible disease type within a few seconds of image acquisition.

Keywords. Mobile standalone application, crop disease, texture, shape, SVM

1 Introduction

Intelligent Decision Support Systems (DSSs) for crop monitoring have been developed worldwide for efficient crop disease and pest management [1, 2]. The systems employ an integrated approach and provide an interactive interface for users to request information to help with crop management. The outputs of the system help to estimate disease risks and provide options for disease control [3]. Many previous studies [4, 5] have proposed frameworks that integrate imaging techniques as a part of the system based on recent cloud technology for preliminary screening of visible diseases in the field. Users can access crop information or request basic disease detection from a server back to their smartphone. In our work, we propose a lightweight but efficient mobile automated system based on imaging techniques and show the potential of the standalone version for real-time diagnosis of crop diseases. This alternative to using a client-server interaction eliminates the delay and transfer time required, especially when users are uploading image data.

In a general framework, the basic imaging procedure includes pre-processing to enhance the image quality and consistency, leaf and disease segmentation, feature extraction and disease classification. Leaf segmentation is essential especially when a

cluttered background is captured. Sophisticated techniques have been developed for leaf segmentation [6]. Our system leverages the programmable smartphone device to automatically eliminate background. For feature extraction, previous studies have demonstrated that the combination of different features, particularly texture, colour and shape features, effectively represent disease characteristics. Work on fusion of features from images acquired under controlled conditions has been undertaken [5, 7, 8]. Our system analytically selects nine sub-features from the three feature sets which are low-complexity and have the potential to characterise disease patterns. Ultimately, our system can provide users with information on whether crops are infected (disease/non-disease) and can also provide information on the possible disease to support crop management.

2 System Overview

Our lightweight mobile application system is illustrated in Fig. 1. The application enables users to capture and/or upload an image (see Fig. 2(a)). If the user selects to upload a photo, only one main leaf is required for the automated classification. The user is required to semi-manually prepare the leaf image. Firstly, the user can call *Auto-rotate* to instigate an automated rotation system to adjust the leaf horizontally; otherwise, the user can manually rotate the leaf clockwise (CW) or counterclockwise (CCW). Then the image background can be removed using a cropping function. If the user has a leaf sample in hand, they can capture the leaf image in the overlay template displayed in a capture screen (see Fig. 2(g)). This approach enables the leaf to have consistent orientation and for the background to be removed automatically. The user can also alter the leaf shape template to correspond to different plants. Once the user calls the *Identify* function, a leaf will be processed for three main feature sets, including colour, texture and shape features. For colour and texture features, the *Cr* colour component (from YCbCr) is used. The co-occurrence matrix is constructed from this component and then three main features - local homogeneity, correlation and entropy of the leaf texture - are computed from the matrix; a global contrast is also included as another texture feature. The same colour component is extracted for basic statistical information, such as mean, standard deviation and skewness. For disease shapes, disease regions are segmented from a leaf using binary thresholding. The segmented disease patches are computed for elongatedness and hydraulic radius properties. In total, nine sub-features from three feature sets are generated to form a feature descriptor for training a binary classification model using a Support Vector Machine (SVM-2 in Fig. 1). This model can differentiate disease from non-disease. After the non-disease leaves are filtered out, only disease feature descriptors are learned for our multiclass model (SVM-5 in Fig. 1) to classify five different diseases, including brown rust (BR), Septoria (ST), yellow rust (YR), powdery mildew (PM), and Tan spots (TS). The application can identify whether the leaf is a diseased or non-diseased leaf; if it is a diseased leaf, the system displays the possible disease type in a small caption (see Fig. 2(f) and (h)) as guidance for the user so that they can apply suitable techniques for controlling the disease.

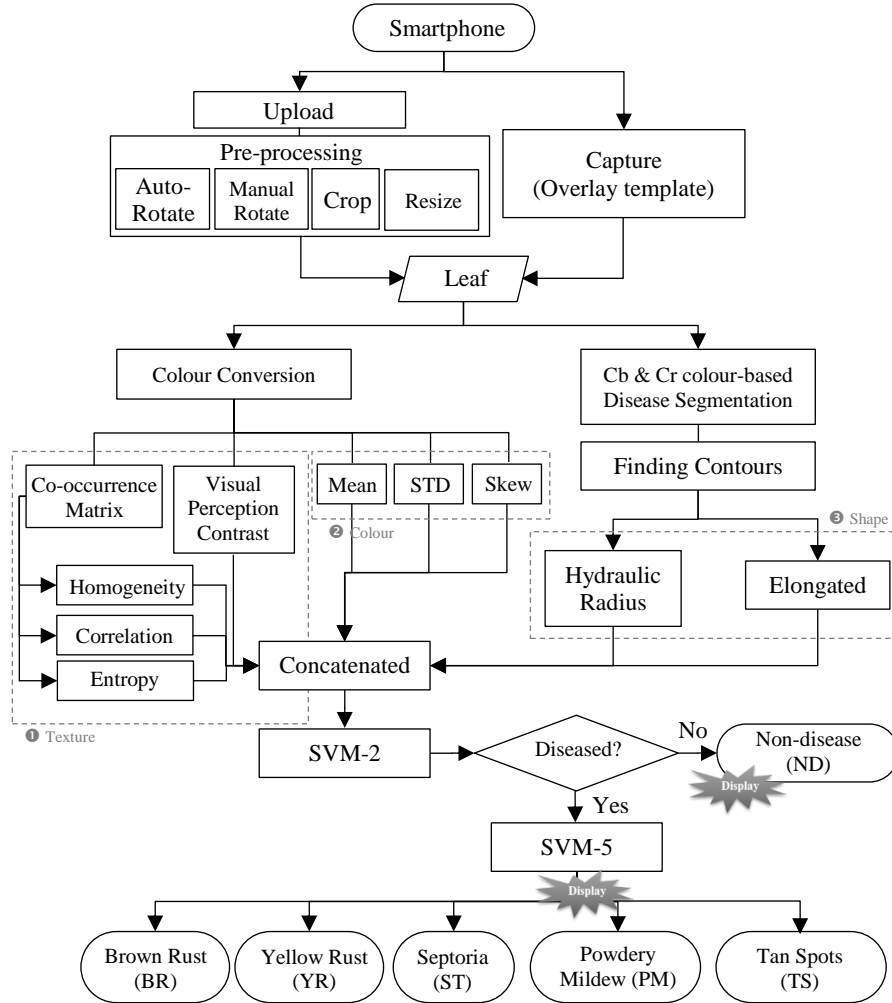


Fig. 1. Mobile System Flow for Monitoring Wheat Disease

3 Semi-automated Pre-processing

Leaf rotation and background removal are important for accurate disease classification. Pre-processing requires user participation only when the user needs to upload an image. Previous research [9] has demonstrated the potential for using a 2-D Discrete Fourier transform (DFT) in detecting global frequency where wheat leaf veins are strongly aligned in a particular direction (corresponding to the orientation of the leaf edge itself). The resulting Fourier spectrum represents a spectrum line that is orthogonal to the leaf orientation. The Fourier transform of an image is implemented in our mobile system; Fig. 2(b) shows a thresholded Fourier spectrum (in white); then the spectrum line is detected using a Hough line transform (HLT) as shown with the red-

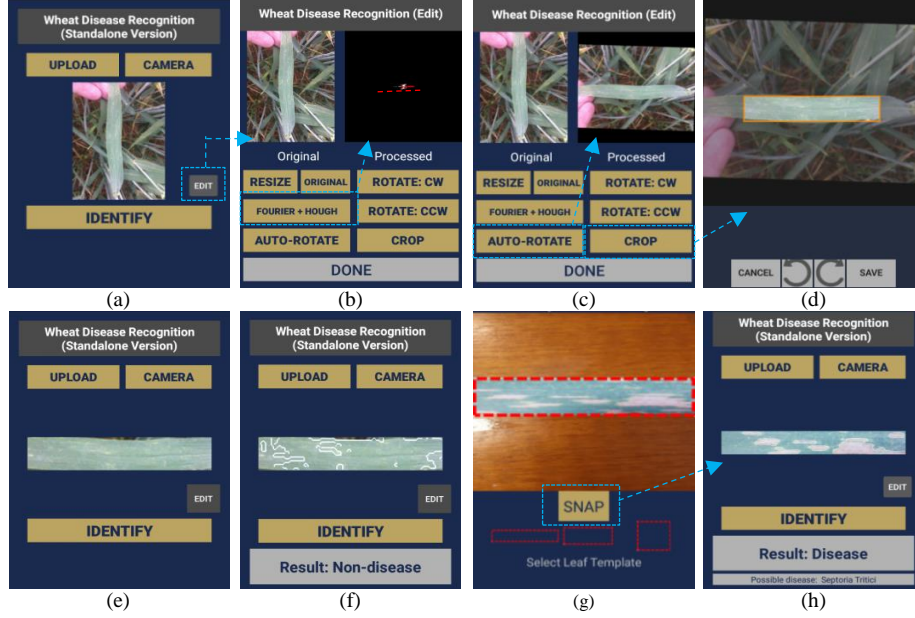


Fig. 2. Targeted leaf acquisition (a) upload an image (b) find leaf orientation using Fourier transform (DFT) and Hough line transform (HLT) (c) rotate leaf after auto-rotating using DFT and HLT (d) manual cropping (e) cropped targeted leaf in the main menu (f) identified leaf type (g) targeted leaf captured in an overlay template (h) identified leaf type

dotted line. The leaf is rotated automatically based on the angle of the strongest detected line from the HLT (Fig. 2(c)). In the case where DFT and HLT cannot perform well for leaf rotation, we provide a manual clockwise (CW) and counterclockwise (CCW) rotation menu for the user. Finally, background removal can be performed by using a *Crop* interface that enables the user to crop the leaf area (Fig. 2(d)) and save it back to the main menu (Fig. 2(e)).

4 Feature Extraction

The extracted leaf region from pre-processing (or from a captured image) is considered for feature extraction. Fig. 3 shows six different disease patches each from a different wheat leaf. It can be seen that non-disease has a homogeneous greenish texture, whereas areas of disease are inhomogeneous in colour distribution and shape. The three main feature sets include texture, colour and shape features which have been demonstrated to best-represent disease patterns [7, 8]. Theoretically, non-disease can be filtered out from the group using texture and colour characteristics. Shape can be used to distinguish striped disease (YR) and the random shape of Septoria (ST) compared to spotty diseases (BR, TS and PM). A number of features were gathered with the choice of features motivated by previous work [5–8]. Nine sub-features which will be described in the following sections were selected based on computational efficiency and individual feature accuracies.

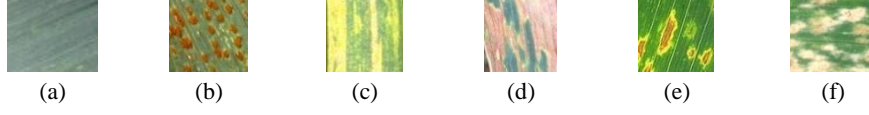


Fig. 3. Different disease patches (a) Non-disease (ND) (b) Brown rust (BR) (c) Yellow rust or stripe rust (YR) (d) Septoria (ST) (e) Tan spots (TS) (f) Powdery mildew (PM)

4.1 Texture Features

The co-occurrence matrix is commonly used to represent spatial relationships of pairs of intensity levels in an image (such as levels i and j that are separated by distance d and direction θ) [10]. We select three main textural features that are derived from the co-occurrence matrix: local homogeneity, a local linear-dependency measure using correlation, and entropy (which represents pattern complexity).

Another textural feature proposed by Tamura et al. [11] is a contrast feature that is based on human visual perception. Tamura contrast considers a global variation of intensity levels and polarization of distribution of high and low intensity areas in an image. The mixture of two components, global variation (or standard deviation, σ) and polarization (or kurtosis, α_4) is used to define the Tamura contrast:

$$F_{con} = \sigma/(\alpha_4)^{1/4} \quad (1)$$

4.2 Colour Features

Basic measures of statistical information in an image are derived from a histogram of intensity values; these are straightforward to compute and include mean (μ), standard deviation (σ) and skewness. Skewness measures asymmetry of the intensity histogram, which can be negative or positive depending on whether the distribution “leans to the left” or “to the right” of the mean value, respectively. Skewness (F_{skew}) can be computed as

$$F_{skew} = \sigma^{-3} \sum_{i=1}^N (i - \mu)^3 h(i) \quad (2)$$

where $h(i)$ is a histogram of intensity i . For both texture and colour features, we empirically selected the Cr colour component (from YCbCr) as a colour representation for the application.

4.3 Shape Features

Disease segmentation is required to extract the disease areas from the main leaf. Otsu thresholding is applied to Cb and Cr colour components of the leaf image, as it has been shown in previous work that these combined colour components for disease segmentation are robust to different lighting conditions [9]. Then the contours of individual disease patches are extracted based on border following algorithms to find connected components of a binary disease image [12]. Fig. 2(f) and (h) show the detected contours (white circle) of non-disease and Septoria diseased leaves. Each con-

tour is processed for two shape features, elongatedness and hydraulic radius. Elongatedness of a diseased patch is based on the ratio of its principal major axis (L_m) and principal minor axis (L_n) and is useful in separating striped disease, such as yellow rust (YR), from other spotty diseases (BR, PM, TS). If a leaf contains N diseased patches, the average value of the N elongatedness values is used to represent the degree of elongatedness degree of that disease:

$$F_{Elong} = \frac{1}{N} \sum_{i=1}^N \frac{L_{m,i}}{L_{n,i}} \quad (3)$$

Hydraulic radius (F_{HR}) is used to measure the complexity of a diseased shape and is a ratio of perimeter (P_d) to disease area (A_d).

$$F_{HR} = \frac{1}{N} \sum_{i=1}^N \frac{A_{d,i}}{P_{d,i}} \quad (4)$$

5 SVM Classification

Nine sub-features (4 textures, 3 colours, and 2 shapes) are combined as a feature descriptor of a leaf and are then used for training a binary classification model (disease/non-disease) using a Support Vector Machine (SVM-2) to filter out non-disease leaves. SVM is applied in the lightweight application as it has been shown to obtain high accuracy and good generalization with economical computation [13]. After the leaf is classified as a diseased leaf, the same descriptors are then used for training a multiclass classification using SVM-5 to provide the user with the possible disease presented in the leaf (a linear kernel is selected empirically for SVM-2 and SVM-5).

6 Experimentation and Results

Our diseased and non-diseased leaf datasets are acquired from various sources including high quality images from the Food and Environment Research Agency (UK FERA) [2], from photographs taken of crops at a farm near Coleraine, Northern Ireland, and from a range of internet sources. The images vary in quality, size and lighting conditions. In preparation, the targeted leaf in each image was segmented manually from the background, as only one main leaf in an image is used for model learning. The proposed system is implemented on a Samsung Galaxy A3 smartphone using OpenCV4Android and libSVM on Android [13]. A total of 610 leaf images are used for learning in the smartphone system, using a 5-fold cross-validation scheme.

The results in Table 1 show the binary classification performance in identifying diseased and non-diseased leaves. The accuracy in detecting disease has high precision at around 92% and recall at approximately 93%, corresponding values for non-disease are approximately 65% and 61%, respectively. One possible impact on the performance of non-disease detection is due to the imbalanced nature of the data between non-disease images (~100) and disease images (~500). The performance of the classifier is much lower on the class with the small sample size (non-disease class). Although techniques for learning from imbalanced data can be employed, such as synthetic data generation and subsampling methods [14], because of our small dataset,

only the original images are used in the experimentation. In addition, there is considerable variation of image quality and size between the datasets used. Overall, disease/non-disease classification accuracy is approximately 88%.


Although some disease types are obvious in their visual appearance, in practice the uncontrolled images have a high degree of variation, which is important for accurately computing colour distribution, overall textures or shapes. Our experimental results based on nine lightweight features and an SVM linear kernel can detect five different diseases, as exhibited in the confusion matrix in Fig. 4(a), which shows the true positive rate of disease identification. Brown rust and Septoria show the highest true positive rates. Fig. 4(b)-(e) shows examples of diseased and non-disease leaves. The non-disease, powdery mildew and Septoria diseased leaves in Fig. 4(b), (c) and (d), respectively are correctly identified. The detected diseased regions are shown with white boundaries. Yellow rust and tan spots have a very low true positive rate, and both tend to be incorrectly classified as Septoria disease. All three diseases including yellow rust, tan spots and Septoria diseases have similar colour distributions, which are mainly yellow and brown. Also, when the severity of disease is high, the disease patches tend to connect to each other, which impacts negatively on the disease segmentation. When all the diseased regions connects, this can result in a random shape similar to Septoria disease, as shown in Fig. 4(e); hence yellow rust is frequently misclassified as Septoria.

	#Image	Precision (%)	Recall (%)	Avg. Acc. (%)
Disease (D)	505	91.98	93.07	87.54
Non-disease (ND)	105	64.65	60.95	

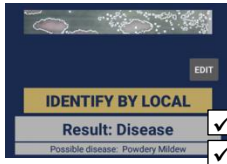
Table 1. Performance measure of Disease vs. Non-disease plant patterns

		Predicted				
		BR	YR	ST	TS	PM
Actual	BR	70.00%	6.67%	20.00%	3.33%	0.00%
	YR	8.28%	18.62%	71.03%	0.69%	1.38%
	ST	0.53%	8.42%	85.26%	1.58%	4.21%
	TS	2.86%	1.43%	80.00%	10.00%	5.71%
	PM	4.29%	1.43%	38.57%	0.00%	55.71%

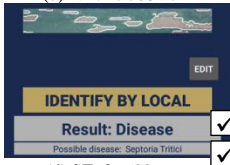
(a) Confusion Matrix



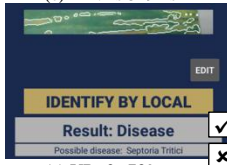
(b) ND: 2s 708ms



(c) PM: 2s 519ms



(d) ST: 3s 699ms



(e) YR: 2s 739ms

Fig. 4. (a) Confusion matrix and % true positive rate for each of the disease (b)-(e) Screen display results of different disease types (Non-disease (ND), Powdery Mildew (PM), Septoria (ST), and Yellow rust (YR)) from Samsung Galaxy A3 with corresponding response time of the image size 2448 x 326 pixels.

We have demonstrated an independent system for lightweight mobile plant disease monitoring that requires low computational effort, taking less than 4 seconds to detect each disease. The user is simply required to crop an image to select a wheat leaf and the image is processed locally, whereas for the other similar work [5], the user is re-

quired to select a particular diseased region in a leaf and the cropped image is sent for processing at the remote server. In addition, non-disease leaf identification which is important for the farmers' early decision in crop management is considered in our application but is omitted in most previous work [5, 7, 8]. The presence of disease is detected with a promising accuracy in our work, although there are a number of false positives (a non-disease detected as a disease). In this stage of our work, for any particular disease detected, the user is required to seek expert advice. Our application size is less than 5MB, but the OpenCV manager application (21MB) is also required.

The system developed considers disease shape, overall texture and colour information inside the leaves regardless of the type or the shape of the leaf itself. Whilst the application is currently targeted at wheat diseases, it could be extended to other plants in the large-scale with similar disease characteristics and integrated with an intelligent decision support system to improve accuracy in classifying a particular disease.

Acknowledgement: I would like to thank Dr. David Gibson, University of Bristol who provided the FERA labelled images and the IU-ATC (EPSRC) for partially funding the project.

References

1. Gent, D.H., De Wolf, E., Pethybridge, S.J.: Perceptions of risk, risk aversion, and barriers to adoption of decision support systems and integrated pest management: an introduction. *The American Phytopathological Society*. 101, 640–643 (2013).
2. FERA: UK Food and Environment Research Agency, <http://www.cropmonitor.co.uk/>.
3. Lucas, J.A.: *Plant Pathology and Plant Pathogens*. Blackwell Science (1988).
4. Prasad, S., Peddoju, S.K., Ghosh, D.: AgroMobile: A Cloud-Based Framework for Agriculturists on Mobile Platform. *International Journal of Advanced Science and Technology*. 59, 41–52 (2013).
5. Xia, Y.: Intelligent Diagnose System of Wheat Diseases Based on Android Phone. *Journal of Information and Computational Science*. 12, 6845–6852 (2015).
6. Gibson, D., Burghardt, T., Campbell, N., Canagarajah, N.: Towards Automating Visual In-field Monitoring of Crop Health. *IEEE International Conference on Image Processing*. pp. 3906–3910 (2015).
7. Tian, Y., Zhao, C., Lu, S., Guo, X.: SVM-based Multiple Classifier System for Recognition of Wheat Leaf Diseases. *Conference on Dependable Computing*. pp. 2–6, Yichang, China (2010).
8. Sarayloo, Z., Asemani, D.: Designing a classifier for automatic detection of fungal diseases in wheat plant. *23rd Iranian Conference on Electrical Engineering*. pp. 1193–1197 (2015).
9. Siricharoen, P., Scotney, B., Morrow, P., Parr, G.: Automated Wheat Disease Classification Under Controlled and Uncontrolled Image Acquisition. *International Conference on Image Analysis and Recognition*. pp. 456–464 (2015).
10. Haralick, R.M., Shanmugam, K., Its' shak, D.: Textural Features for Image Classification. *Systems, Man and Cybernetics, IEEE Transactions on*. 3, 613–621 (1973).
11. Tamura, H., Mori, S., Yamawaki, T.: Textural Features Corresponding to Visual Perception. *Systems, Man and Cybernetics, IEEE Transactions on*. 8, 460–473 (1978).
12. Suzuki, S., Abe, K.: Topological Structural Analysis of Digitized Binary Images by Border Following. *Computer Vision, Graphics, and Image Processing*. 30, 32–46 (1985).
13. Chih-Wei Hsu, Chih-Chung Chang, and C.-J.L.: A Practical Guide to Support Vector Classification, <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
14. He, H., Garcia, E.A.: Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*. 21, 1263–1284 (2009).